ORSETT PSYCHOLOGICAL REVIEW

No.9 - MARCH 2003

A JOURNAL ABOUT THE PSYCHOLOGY OF EVERYDAY LIFE

ISSN: 1474-0311

Orsett Psychological Services
PO Box 179
Grays
Essex
RM16 3EW

CONTENTS

LINKED ARTICLES - INTELLIGENCE

This is a series of articles about the topic of intelligence. The aim of the series is to look at intelligence in different situations, including humans, non-human animals, and machines.

No.1 - Brewer, K (2002) Non-Human Animal Intelligence Orsett Psychological Review September, 2-13

No.2 - Brewer, K & Reinicke, H (2002) A Brief Survey of Human Intelligence Orsett Psychological Review December, 8-11

No.3 - Human Cognition and Cognitive Science: A Review of the Limitations of Modelling Human Cognition using Computer Technology - Part 6: Intelligence

INTRODUCTION

The question of whether or not it is possible to create an intelligent machine still remains as a key philosophical debate in AI. The fact that certain computer programs can perform extremely complex tasks involving processing that exhibits seemingly "intelligent" behaviour is not at question here. Some notable examples of very advanced AI systems include;

- The "Deep Blue" chess program that beats chess masters (IBM 2003).

- The GASOIL expert system using so-called "decision-tree" machine learning methods could outperform human experts in design gas-oil separation systems for oil platforms (Russell and Norvig 1995 p539).

- Learning to fly an aircraft in flight simulator from observing human pilots doing the task. Sammut et al (1992) describe how machine learning techniques were used to devise rules for flying the aircraft. These rules were used to create a computer program that was capable of flying a simulated aircraft better than some of the human instructors (Russell and Norvig 1995 p539).

But if one poses the question - "are these systems intelligent?" - general opinion would suggest that they are not. How is it that computers can exhibit intelligent behaviour, and yet not be intelligent? The lack of any sense of self-awareness, of "ego", of

consciousness appears to be of fundamental importance. These systems, as remarkable as they are, are still essentially following instructions (algorithms and heuristics, sometimes in conjunction with a particular architecture as in the case of neural network systems) to produce behaviours. They have no notion of themselves.

In this article, we examine the ways in which computers can perform various aspects of intelligent behaviour and then we turn to the issue of consciousness in machines. We do not claim to answer the question of whether it will ever be possible to build a conscious, and hence arguably an intelligent machine. But we do suggest that these issues are linked and that it is unlikely that true machine intelligence will be widely recognised without the evidence to show that the same machine exhibits some degree of consciousness or self-awareness. Perhaps the striving for machine consciousness will become the ultimate goal for AI practitioners in the attempt to win the machine intelligence debate.

The rest of this article contains the following sections. Firstly, we list some of the phenomena of intelligence that were introduced in Brewer (2002) and which are used in this article to gauge machine intelligence. We then look briefly at the way in which some of these phenomena are combined to produce cognitive architectures (cognitive models for programming systems that produce intelligence behaviours when they operate). We then describe in more detail how these phenomena are handled in machines to produce intelligent behaviours. Finally, we present some thoughts on machine consciousness.

SOME PHENOMENA OF INTELLIGENCE

Brewer (2002) presents intelligence phenomena that are used to compare animal intelligence with human intelligence. Brewer (2002) states that overall intelligence is composed of separate abilities (so-called "intelligences" - such as perception and symbol manipulation). In animals, not all intelligences are found to be completely developed in any one species. In general, abilities needed to survive in the environment appear to be most highly developed. Although certain animal species may excel at particular intelligences (such as perception), no individual species has the combination of all intelligences in the way that humans do, and so animal intelligence is inferior to human

- intelligence (Brewer and Reinicke 2002). It is instructive to use the "intelligences" given in Brewer (2002) to gauge machine intelligence here. These are

- Sensory abilities (perception)
- Memory
- Learning
- Problem Solving, Planning, Creativity
- Language and Symbol Manipulation
- Self Awareness (awareness of one's own mind)
- Theory of Mind (awareness of the mind of others)
- Culture
- Consciousness

Cognitive scientists, such as Newell, Young and Polk (1993) confirm that many of these phenomena are required within a general theory of cognition. A general (or unified) theory of cognition describes a single system of mechanisms that operate together to produce a full range of cognition (Norman 1991). In particular, Newell et al (1993) suggest that a unified cognitive theory should include such "intelligences" as sensory ability (perception), memory, learning, problem solving, planning, and symbol manipulation (including language manipulation).

Much research has been carried out in computer science to improve and develop the way in which computers perform at these abilities, it is not surprising that certain computer systems are very capable in these areas. However, there are very few cognitive theories to describe the operation of self awareness, theory of mind, interaction of the cognitive entity within its culture, and consciousness. As a result, computers have limited capability in these areas.

COGNITIVE ARCHITECTURE AND MODELS

One of the ground assumptions of cognitive science is "the computational view of the mind" - ie the view that cognition in humans may be likened to information processing in computers in the sense that humans and computers store, retrieve and use information in a similar manner. As computer architectures and AI techniques have become more powerful and flexible, computers themselves have also provided ways of thinking about how humans cognition might operate (Medin, Ross and Markman 2001). For example, consider the close association of cognitive neuroscience and the study of neural networks in AI.

Within cognitive psychology, numerous studies have been carried out to investigate how humans are capable of carrying out the "intelligences" listed above. These studies have led to individual theories that describe how, for example, learning, problem solving, memory might

occur in the brain (or, more generally, in an information processing system). Some of these theories have been used, developed and/or combined to produce computer architectures for modelling aspects of human cognition (1). Some of these cognitive architectures include relatively well known AI "languages" such as Soar and ACT-R. The book by Pew and Mavor (1998) is a good starting point for further reading about these, and other cognitive architectures.

Though space does not permit a detailed examination of any single cognitive architecture per se, for the sake of clarity, a generic "general purpose" cognitive architecture is based upon the "human information processor" as illustrated in Figure 1. This contains the following key components:

- Sensors and Perceptors: Facilities and functions that transform environmental stimuli into internal representations (in computers, these are binary coded representations) which are manipulated by the cognitive processor.

- Working (Short-Term) Memory: This contains information stored temporarily in memory during cognitive processing (see description of memory in subsequent paragraphs).

- Long Term Memory: This contains large amounts of data that are stored over long periods of time, and includes the experiential knowledge and "expertise" of the system.

- Cognitive Processor: This incorporates a range of functions that carry out information processing - eg planning, learning, and decision making.

STIMULI
auditory info
visual info etc

COGNITION
* planning
* learning
* decision-making

RESPONSES
actions
speech etc

→ ↑ ↓ →

SENSING AND PERCEPTION → ← WORKING MEMORY → OVERT BEHAVIOUR
* vision
* hearing
* perception

↑ ↓
LONG TERM MEMORY

Human information processor

Figure 1 - Generalised cognitive architecture based on Pew and Mavor (1998).

- Overt Behaviour Effectors: These include functions and components that carry out physical actions within the environment that are the result of cognitive processing of incoming data (from the sensors and perceptors). These might include speech generation components, motor components for movement.

COMPUTER MODELS AND THE "INTELLIGENCES"

An overview of the way in which computer models and cognitive architectures handle each of the intelligence types listed above and in Brewer (2002) is given below.

Sensory abilities (perception)

As stated in Brewer (2002), highly refined perceptual abilities in an organism (or machine) alone are not indicative of intelligence, but intelligent systems generally require some perceptual ability to allow them to interact with the environment.

Perception in animals and machines can be highly skilled and tailored to specific requirements or activities. For example, bats use highly specialised echo-location akin to sonar to navigate while flying and to search for prey. Similarly, in ships and aircraft, specialised computer systems incorporate sonar and radars for navigation (however, we do not claim that these machine systems are cognitive!). Where machine perception is linked to pattern matching, such as the use of bar-code readers to identify cost tags on goods bought in shops, these systems are not "intelligent" at all.

Where cognitive machine systems incorporate perceptual functionality, eg in certain "autonomous" robot systems, such as self-driving cars, perceptual abilities provide some of the input necessary for perceptual processing that feeds data into a decision making or cognitive system.

Using the "self driven car" example, the extraction of image data from camera (video) images - such as the recognition of shapes and 3D cues from what are essentially 2D images is certainly a very skilled process (see Medin, Ross and Markman 2001 p80). But the act of cognition (taking these images and controlling the vehicle as a result) is much wider than pattern recognition alone, and involves using image cues to make decisions about manipulating the steering, accelerator and brake functions.

Irrespective of whether machine perception is linked to cognitive or to traditional non-cognitive computer

processing systems, current state-of-the-art computer perception systems are very specialised and in many cases are capable of detecting features imperceptible to humans.

This is the case for infra-red, ultra-violet, sonar, radar, and the machine detection of other radio frequency signals, and for detecting vibrations (seismic disturbances), pressure differences (eg weather), bio-chemical contaminants amongst many other environmental factors. Whereas machine perception excels in carrying out very specialised tasks such as these, human perception tends to be applicable to carrying out a much more general range of tasks and activities. Thus, as an overall, general purpose visual perception system capable of handling tasks such as face, text, and shape recognition, human visual cognition still exceeds machine visual cognition.

Memory

Memory plays a fundamental role in cognition. In humans, some of the functions that memory serves includes the following:

- It stores knowledge in such a way that we can retain a limited number of facts and derive other facts from them by inference when needed.

- It relates new events and facts to prior knowledge in order to understand them.

- It delivers relevant knowledge when needed.

In both humans and computers, two types of memory system are recognised. These include short term memory and long term memory. Short term memory is sometimes called working memory - we will use the term "working memory" in the remainder of this article.

Working memory holds important information in a readily available state so that it can be used for reasoning or comprehension tasks. Working memory is also used to retain information prior to storage in long term memory. Information in working memory can also be augmented with information retrieved from long term memory during problem solving.

In humans, working memory is limited in capacity (2). On the contrary, in computer systems, working memory is much more extensive. Computer working memory may be thought of as non-persistent memory that computer programs use while they are running. This memory disappears when the computer is switched off (ie it is

not retained on a physical medium such as a hard disk or CD ROM), but while running, this memory is stored in RAM, in the computer registers and on the program stack (see Norton and Goodman 1999 p160 and p293). The working memory limitations of modern computers far exceed many program requirements and at the time of writing are in the order of tens of mega bytes. Within cognitive "programs" themselves - programs created using cognitive architectures/languages - working memory describes the current program state, with its relevant program variables and their associated values representing the results of intermediate computations.

Information may pass from working memory into long term memory through a number of possible mechanisms which we cannot elaborate on here (see Medin, Ross and Markman 2001 for more information). In humans, long term memory describes what is essentially persistent memory; except for the fact that this memory is not always infallible. The inability to recall information from memory (forgetting) has been attributed to a number of factors, such as memory decay, interference and overwriting.

However, it seems that the quality of long term memory depends upon how the information has been encoded and the context at learning, as well as the context at the time of retrieval. For computers, long term memory relates to information stored persistently on the computers' hard disks, or on some other storage medium (floppy disk, tape, CD ROM etc). In general, unless there is a fault with the storage medium itself, or corruption of the medium (eg due to magnetism), computer long term memory is infallible. Within cognitive programs themselves, long term memory describes the persistent knowledge encapsulated within some kind of knowledge representation language (eg production rules in the Soar cognitive architecture).

Learning

Learning is an important feature of intelligent systems. Learning often implies "profiting by experience" and may be manifested as an improvement in performance or by a greater adaptation of a cognitive entity to its environment. Learning can also relate to the storage of new information in memory which may or may not result in an observable change in performance. Learning in humans and animals is discussed briefly in Brewer (2002), and the cognitive basis for learning is explained in detail in Medin, Ross and Markman (2001). We have also touched on computer learning in a previous article in this series which takes a slightly different slant to the description below - ie previously we have focussed on learning and language theory, and learning in neural networks (see

Allsopp and Brewer 2001a).

In recent years, there has been a proliferation of so-called machine learning techniques. Machine learning describes computer paradigms and algorithms that allow program behaviour to be altered by experience, and/or allow new information about the environment to be stored in memory. A few of these techniques include neural networks, decision trees, algorithmic and probabilistic techniques such as Bayesian learning and genetic algorithms. See Russell and Norvig (1995) (summarised in Kalus et al 2003 p25) for an overview of these and other AI learning methods.

Many cognitive architectures also contain in-built learning capability. For example, in the Soar cognitive architecture, this learning ability is termed "chunking" (Laird, Newell and Rosenbloom 1987 p35). Soar programs use production rules to represent long term memory or knowledge. During "chunking" Soar creates new production rules to describe "short cuts" that the architecture may have found during problem solving after encountering a so-called "impasse".

Impasses occur when Soar knowledge (represented in its production rules) becomes exhausted while the program operates to achieve its eventual goal state. During an impasse, the Soar architecture uses default knowledge (within default rules) to search for a solution to the unresolved sub-state that it creates. If it manages to resolve the impasse, the previously unresolved sub-state is captured in a new production rule which encodes the solution found for use next time it is encountered. (Also see below for a discussion about problem solving in the Soar architecture).

Problem Solving, Planning, Creativity

In this section, we will concentrate on problem solving only. Planning may be considered as being related to a particular type of problem solving ie one in which the goal is to derive a future course of action. Medin et al (2001) also argue that creativity may be linked to problem solving.

A problem may be characterised by 4 aspects:

- Goal: A state of knowledge toward which problem solving is directed.

- Givens: Objects, conditions, constraints associated with the problem.

- Means of transforming conditions: Ways of changing initial states to goal states.

- Obstacles: Missing knowledge or "unknowns" that must be "filled in or elicited during problem solving in order to attain the goal.

Thus:

> Problem solving is taking place if a person is (1) trying to attain a goal, (2) starting from some set of conditions, the givens, (3) with some means of transforming these conditions, but (4) with no immediately available knowledge of a solution (Medin et al 2001).

Newell and Simon's (1972) influential view of problem solving sees problem solvers (people or machines) as information processing systems with particular characteristics, eg in which processing occurs in a serial fashion, in which there is a limit to working memory, and in which long term memory is unbounded.

They suggested that problem solving occurs within a cognitive representation called the problem space: this is the problem solver's internal representation of the problem, consisting of states and operators. A problem state consists of the knowledge available to the problem solver about the problem "environment" at a particular time. For most problems, many problem states must be traversed to get from the initial problem state to the finished goal state.

Operators describe the cognitive operations that occur when moving from one problem space to another - only one operator can be used at a time to move problem solving on to the next state. This view of problem solving is termed problem solving by "representation and search" and is has been built into the Soar cognitive architecture (Laird, Newell and Rosenbloom 1987).

In order to move from initial to goal state, some kind of "look ahead" state search is needed to determine which might be a suitable problem space to progress to given the current state. In problem solving by "representation and search", heuristics are used to guide this search. Heuristics are general rules and guidelines that do not guarantee a solution, but allow a good chance of finding a solution without exhaustively searching through all possibilities.

A number of simple search heuristics include "hill climbing" and "means ends analysis" (see Medin et al 2001). Heuristics like these are used in many AI problem solving and planning systems, and may be readily programmed in Soar using few production rules owing to Soar's predisposition to this kind of problem solving strategy.

Although this view of problem solving has greatly influenced AI, other AI techniques are also available for problem solving and these have particular advantages in situations where previous experiences (situations) or when so-called "expert knowledge" is a vital ingredient in attaining the solution. Two examples are briefly described here to illustrate this:

Case Based Reasoning (CBR):

CBR solves new problems by adapting previously successful solutions to similar problems (Kruusmaa and Willemson 2002). Central to the CBR approach is the notion of a "case", which may be thought of as an abstraction of a series of events (solved or unsolved problems) or a set of data. Past experiences are stored in a case base, which resembles a database.

Cases are indexed, and when a new problem occurs, the indexes are extracted from its features and used to match cases (and their associated solutions) in the case-base. If there is no close match, a modified solution may have to be derived before it can be used in problem solving. If the solution is a good one, it is stored in the case base for future reference.

CBR involves four phases (taken from Kirsopp and Shepperd 2002):

- the retrieval of similar cases.
- the reuse of retrieved cases to find a solution to a problem.
- where necessary, the revision of the proposed solution if no close match is found.
- the retention of the solution to form a new case.

Expert systems:

Where expert knowledge of a particular domain is both well understood and subject matter experts are available for knowledge capture, eg for certain medical diagnosis tasks, expert systems may be the most appropriate technology for building problem solvers.

Expert systems attempt to simulate the reasoning of experts. They comprise of a rule base or rule bases (which contain the expert knowledge of their domain) and an inference engine (a software component which allows the rule base to be searched for possible answers to input queries).

Expert systems are by nature constructed for very specific tasks, eg the MYCIN expert system for medical diagnosis (Shortliffe 1976).

Despite their "intelligent behaviour", both CBR and

expert systems are not cognitive systems and have no other intelligent
ability other their problem solving abilities that only operate within the very specific domain for which they were intended.

Language Manipulation

In this section we briefly examine language manipulation in computer systems. To an extent, the review of "problem solving" above defines one view of the cognitive basis of symbol manipulation in AI systems. For further reading, see Newell (1990); Newell, Young and Polk (1993).

Cognitive studies of language manipulation tends to focus on three main aspects of language:

- The sounds of the language itself - language utterances - termed "phonemes".
- The meaning of units of language - termed "morphemes".
- The grammatical structure of the language - termed language "syntax".

Although there are still problems to be overcome, specialised computer systems exist which have the capability to manipulate natural language in one, two or all of these areas (though there are very few that have skills in all three of these areas).

Speech recognition and speech generation:

Speech recognition is a task which cannot be "programmed" into the computer, or coded using an AI program or expert system. Speech recognition is a "fuzzy" task - there are no absolute parameters to describe the pronunciation of a word, as regional and international accents vary.

Neural network technologies have therefore been applied to solve this problem (neural networks are generally good in fuzzy domains where plenty of network training data is available). Problems of speech recognition have been resolved to the extent that it is now possible to buy commercial software applications with a speech recognition component - eg word processors that operate by dictation.

The neural network embedded in this kind of software has to be trained, and hence the software is not always accurate. For example, the software cannot recognise new words, so the user must select from a list of possible contenders, or must type in the new word. But accuracy at

recognising spoken words improves with use, and a well-trained network is very effective at understanding the spoken words of its trainer.

In most modern speech generation systems, digitised words, phrases and sentences are used, which are joined together to make utterances. In many speech generation systems (eg automated telephone system), there is usually little, or no understanding of the meaning of the utterances made. Often, entire digitised sentence fragments are selected from a program where the "changeable" parts of the sentence are missing. For example, when calling an automated banking telephone service for an account balance, the balance of the account is missing from the response phrase - when the balance is found, it is converted into a spoken number, and slotted into the answer response. In earlier speech generation systems, synthetically generate voice sounds were common - but these sound more unnatural.

Understanding language meaning (semantics) and syntax (grammar):

Understanding the syntax of natural language once it is "tokenised" (ie represented as symbols, such as ASCII text) inside the computer is a complex process which involves a number of tasks. These include: decomposing each sentence into its phrase structure (ie identifying the noun and verb phrases within the sentence, decomposing these into constituent nouns, verbs, articles, adjectives etc), and assessing the "deep structure" of the sentence, eg in a question phrase in English, the position of the verb and object can be reversed (see Chomsky 1957, 1981, 1988).

Syntactic analysis occurs one word at a time and syntactic structures are assigned to words as they are encountered. Where many syntactic structures may be assigned to a word (eg "read" could be a noun, verb etc), heuristics are used to select a candidate. If, on further analysis of the sentence, this selection is incorrect, the sentence must be reanalysed (parsed) to elicit its correct structure (Medin, Ross, and Markman 2001 p355).

Although language parsers are reasonably effective at determining the underlying syntax of sentences (all modern word processors have spelling and grammar checkers), a true interpretation of a sentence cannot be made from syntactic information alone. Cognitive psychologists still do not completely understand how the process of language comprehension occurs in humans.

When understanding language, humans tend to make inferences about what is being said and thus fill in details that are not stated directly. Different theories

have been put forward to describe the nature of these inferences - at present, the debate over which types of inferences are made has not been resolved (see Graesser, Millis and Zwaan 1997; McKoon and Ratcliff 1992; and Graesser, Singer and Trabasso 1994). Thus AI researchers have struggled to model language semantic processors as they are still unclear how to integrate "background knowledge" with new information described within the sentences being analysed.

Self Awareness and Theory of Mind

Self awareness - or the awareness of one's own mind - overlaps to an extent with consciousness, which is covered in more detail below. To date, a cognitive model with true self awareness does not exist, although it is possible to build "self preservation instincts" into cognitive models using behavioural moderators as described in Allsopp and Brewer (2001b). The theme of self awareness in machines will be dealt with in more detail in a forthcoming article in this series which covers self actualisation, leadership and motivation (personality) in cognitive computer models.

Theory of mind - or the awareness of the mind of others - is still rather a distant goal of AI given that machines with self-awareness are still unavailable. However, current work at MIT on the "Kismet" robot shows that it is possible to develop a cognitive model which exhibits social behaviours relevant to specific social contexts.

In the Kismet project, this social setting is that of the carer/child relationship, where the robot takes the "child" role (MIT 2003a).

Kismet makes use of so-called human social protocol" in its interactions. For example, when the "carer" approaches the robot too closely, Kismet draws back (by craning its neck backwards) to signal this. Similarly, if the carer is too far away to be seen clearly (as determined by the resolution of the robot's cameras), Kismet calls the carer to come closer. The robot is capable of conducting conversations with its carer, and capable of expressing emotional reactions to external and internal stimuli.

The architecture of Kismet is a refinement of the general cognitive processor described in Figure 1, but in essence it contains all of the salient features. The perceptual and attentive system is highly refined, as well as its "effector" system, and Kismet incorporates a "motivation" system which defines its homeostatic needs and its emotional responses.

Kismet research is a step towards a machine that interacts in human social contexts. But, as stated in the Kismet website:

> developing Kismet raises questions about: self identity, theory of mind, autobiographical memory, recognition of self, other, and conspecifics, social learning (esp. imitation), intentionality, emotion, empathy, personality, friendship, ethics
> (MIT, 2003a).

In terms of robotics and cognitive modelling, these questions currently remain unanswered.

Culture

Cultural factors have been incorporated within cognitive models used in certain types of computer simulations (Lannon, Klein and Timian 2001)). The modelling of culture within cognitive models has been already been summarised in Allsopp and Brewer (2002). In addition, the modelling of social groups of cognitive entities has been attempted by MIT (cf the "ants" micro-robot program; MIT 2003b).

However, without some degree of self-awareness and theory of mind, it is unlikely that cognitive models will accurately portray or model cultural effects. Until then, cognitive models will continue to rely on general heuristics (or "rules of thumb") to define their cultural leanings, and these heuristics are likely to remain limited in scope and stigmatic.

Consciousness

Cognitive psychologists and biologists cannot yet explain how consciousness occurs. Theories have been suggested which are at best incomplete, and which appear to be based more upon philosophical argument than upon scientific fact. These theories span the extremes between the entirely materialistic view of consciousness on the one hand, to the dualistic, "spiritual" view on the other. These opposing positions are explained thus:

Materialistic view:

This view of consciousness assumes that consciousness "resides" in the brain and is entirely a product of brain activity. It suggests that the brain, due to its structure and activity, fabricates the state of consciousness (Searle 1992):

> the truly incredible thing is that a collection of simple cells can lead to thought, action and consciousness (Russell and Norvig 1995).

Here there is no place for a spiritual dimension (call it what you will), consciousness is not seen as an ethereal process that resides outside the brain at any time. The materialistic view is expounded in detail in Dennet (1993), who draws upon neurology and computer science to argue that the activity of the brain as a "massively parallel processor" provides the platform for consciousness.

Dualist view:

So called "mystic" or dualistic beliefs that consciousness is linked in some way to the "spirit" or similar realm beyond the scope of physical science also exist. These notions have root in earlier philosophical traditions, eg Descartes (1637), and are supported by a number of more contemporary academics eg Koestler (1967); Popper and Eccles (1977).

The dualist view of consciousness is also supported by reports of "out of body" (OBE) or near death experiences (NDE). There are a limited number of "scientific" cases of these phenomena: eg Pam Reynolds whose brain was frozen during an operation was clinically dead, yet reported later looking down on her body and could recall details about the operation ("The Day I Died" 2003; BBC Television). Others are sceptical (eg Blackmore 1993).

Dualism relegates consciousness to a place beyond our current scientific reach, beyond experimental analysis and any hope of being modelled in the computer.

Whichever view of consciousness you hold to, intractable problems await. Blackmore (2001) states:

> If you think that the mind is something different from matter then you have problems (as Descartes did) with how these two worlds interact. If you deny a separate mind and stick to only physical brains and neurons, then you deny your own subjective experience. If you accept that there is subjective experience, and that there are also physical brains that cause that experience then you have to bridge the "explanatory gap", or what William James called the "fathomless abyss" or the "chasm between the inner and outer worlds" (James, 1890/1983) (Blackmore 2001).

In order to get around these seemingly "intractable problems", some researchers have argued that there is no problem at all - that consciousness will be understood

when we have greater knowledge of how the brain works. Others are convinced that the problem of understanding consciousness is so difficult that it requires a revolution in physics (Penrose 1994).

Still others believe that consciousness will always be beyond our ability to fully understand (McGinn 1999). Indeed there may be problems with our approach to understanding consciousness itself that need to be addressed before further progress can be made (Blackmore 2001).

Despite the fact that our understanding of consciousness is still incomplete, it seems likely that those pursuing the goal of machine consciousness will not succeed if the former materialist theory of consciousness proves to be incorrect. Even if it were correct, not all materialists believe that machine consciousness will ever be possible. Searle (1992) believes that although consciousness is entirely a product of the brain, that it cannot exist anywhere other than in the brain and cannot be artificially created.

On the other hand, there are many cognitive scientists that believe that it sooner or later it will eventually be possible to build a conscious machine - that this is only a matter of time. Igor Aleksander (2001) claims that further development of consciousness in machines depends upon greater understanding of how the machinery of the brain works.

The more knowledge we have, the more we can apply this knowledge to building ever more complex neural network architectures. Aleksander states that machine consciousness will not be like human consciousness, and we should not expect it to be so. He suggests that if we were to build a neural network system complex enough to mimic human brain activity, machine consciousness would emerge as a result.

SUMMARY

We have examined machine intelligence briefly in this report, and have looked at the components of intelligence, ie the "intelligences", and how machines fare in modelling each of these aspects of intelligence. We have introduced the notion of a cognitive architecture as a means of modelling intelligent behaviour in computers.

These cognitive architectures incorporate many of the "intelligences" listed in this report, namely: sensory abilities, memory, learning, problem solving, planning, and symbol manipulation. There are many computer systems that are highly refined in modelling individual aspects or combinations of all of these

intelligences.

However, most cognitive models do not incorporate what we might call the more "abstract" intelligences, such as self-awareness, theory of mind, culture and consciousness. Cognitive theories that describe these abstract intelligences are incomplete, where they exist at all. As a result, machines have very little capability in these areas. For this reason we believe that it is not possible to create a truly intelligent machine at this time.

FOOTNOTES

1. We state that only certain aspects of human cognition are modelled in these architectures because no cognitive model yet exists that incorporates theories that describe how all the "intelligences" operate.

2. Miller (1956) argues that short-term memory can hold between approximately five to nine pieces of information at any one time. While the Working Memory System model (Baddeley and Hitch 1974) has a limited capacity built into it.

REFERENCES

Aleksander, I (2001) How to Build a Mind London: Phoenix

Allsopp, D.J & Brewer, K.M (2001a) Human Cognition and Cognitive Science: A Review of the Limitations of Modelling Human Cognition using Computer Technology - Part 3: Intuition Orsett Psychological Review No.3, September, 2-10

Allsopp, D.J & Brewer, K.M (2001b) Human Cognition and Cognitive Science: A Review of the Limitations of Modelling Human Cognition using Computer Technology - Part 4: Physiological and Environmental Factors Orsett Psychological Review No.4, December, 8-16

Allsopp, D.J & Brewer, K.M (2002) Human Cognition and Cognitive Science: A Review of the Limitations of Modelling Human Cognition using Computer Technology - Part 5: Sociological and Cultural Factors Orsett Psychological Review No.5, March, 12-18

Baddeley, A & Hitch, G (1974) Working Memory. In Bower, G.A (ed) Recent Advances in Learning and Motivation, Vol 8 New York: Academic Press

Blackmore, S (1993) Dying to Live London: Grafton

Blackmore, S (2001) State of the art: consciousness The Psychologist Vol.14, No.10, October, 522-525.

Brewer, K.M (2002) Linked Articles - Intelligence: No.1 - Non-Human Animal Intelligence Orsett Psychological Review No.7, September, 2-13

Brewer, K.M & Reinicke, H.L (2002) Linked Articles - Intelligence: No.2 - A Brief Survey of Human Intelligence Orsett Psychological Review No.8, December, 8-11

Chomsky, N (1957) Syntactic Structures The Hague, Netherlands: Mouton

Chomsky, N (1981) Lectures on Government and Binding Dordrecht, Netherlands: Foris

Chomsky, N (1988) Language and the Problems of Knowledge: The Managua Lectures Cambridge, MA: MIT Press

Dennett, D.C (1993) Consciousness Explained London: Penguin Books

Descartes, R (1637) Discourse on Method Paris

Graesser, A.C; Millis, K.K & Zwaan, R.A (1997) Discourse comprehension Annual Review of Psychology 48, 163-189

Graesser, A.C; Singer, M & Trabasso, T (1994) Constructing inferences during narrative text comprehension Psychological Review 101, 371-395

IBM (2003) Deep Blue Wins. Last accessed March 18, 2003, from http://www.research.ibm.com/deepblue/home/html/b.html

James, W (1983) The Principles of Psychology Cambridge, MA: Harvard University Press (original work published 1890)

Kalus, T; Navarro-Prieto, R; Conaty, G & Allsopp, D.J (2003) Application of Adaptive User Profiling for Reconfigurable Systems IST-2001-34091 SCOUT Work Package 2, Deliverable D2.2 (available from author)

Kirsopp, C & Shepperd, M (2002) Case and Feature Subset Selection in Case-Based Software Project Effort Prediction. In Bramer, M; Preece, A & Coenen, F (eds) Research and Development in Intelligent Systems XIX: Proceedings of the 22nd SGAI International Conference on Knowledge Based Systems and Applied Artificial Intelligence (pp 61-74) Cambridge, UK: Springer Verlag

Koestler, A (1967) The Ghost in the Machine New York: Macmillan

Kruusmaa, M & Willemson, J (2002) Covering the Path Space: A Casebase Analysis for Mobile Robot Path Planning. In Bramer, M; Preece, A & Coenen, F (eds) Research and Development in Intelligent Systems XIX: Proceedings of the 22nd SGAI International Conference on Knowledge Based Systems and Applied Artificial Intelligence (pp 4-16) Cambridge, UK: Springer Verlag

Laird, J.E; Newell, A & Rosenbloom, P.S (1987) Soar: An Architecture for General Intelligence Artificial Intelligence 33, 1-64

Lannon, G.P; Klein, H.A & Timian, D.H (2001) Integrating Cultural Factors into Threat Conceptual Models. In Proceedings of the 10th Conference on Computer Generated Forces and Behaviour Representation. Norfolk, Virginia, USA. 15th-17t May 2001

McGinn, C (1999) The Mysterious Flame: Conscious Minds in a Material World New York: Basic Books

McKoon, G & Ratcliff, R (1992) Inference during reading Psychological Review 99, 440-466

Medin, D.L; Ross, B.H & Markman, A.B (2001) Cognitive Psychology (3rd ed) New York: Harcourt

Miller, G (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information Psychological Review 63, 81-97

MIT (2003a) Kismet. Last accessed on March 18, 2003 from http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html

MIT (2003b) The Ants: A Community of Micro-Robots. Last accessed on March 24, 2003 from http://www.ai.mit.edu/projects/ants

Newell, A (1990) Unified Theories of Cognition Cambridge, MA: Harvard University Press

Newell, A & Simon, H.A (1972) Human Problem-solving Englewood Cliffs, NJ: Prentice-Hall

Newell, A; Young, R & Polk, T (1993) The Approach to Symbols. In Broadbent, D (ed) The Simulation of Human Intelligence Oxford: Blackwell

Norman, D.A (1991) Approaches to the Study of Intelligence Artificial Intelligence 47, 327-346

Norton, P & Goodman, J (1999) Inside the PC: Eighth Edition SAMS

Penrose, R (1994) Shadows of the Mind Oxford: Oxford University Press

Pew, R.W & Mavor, A.S (1998) Modelling Human and Organisational Behaviour: Application to Military Simulations, Panel on Modelling Human Behaviour and Command Decision Making: Representations for Military Simulations, Commission on Behavioural and Social Sciences Education, National Research Council National Academy Press: Washington DC

Popper, K.R & Eccles, J.C (1977) The Self and Its Brain Berlin: Springer-Verlag

Russell, S.J & Norvig, P (1995) Artificial Intelligence, A Modern Approach Prentice-Hall International

Sammut, C; Hurst, S; Kedzier, D & Michie, D (1992) Learning to Fly. In Proceedings of the Ninth International Conference on Machine Learning, Aberdeen Morgan Kaufmann

Searle, J.R (1992) The Rediscovery of the Mind Cambridge, MA: MIT Press

Shortliffe, E.H (1976) Computer-based medical consultations: MYCIN New York: Elsevier

Daniel Allsopp & Kevin Brewer

Article written March 2003

THE BARNUM EFFECT - A SMALL SCALE STUDY

INTRODUCTION

The Barnum effect and Pollyanna mechanisms formed the theoretical basis for this study. The Barnum effect is a belief that a generalised stereotyped statement is an accurate self-description and personally specific. This is due to such statements being derived from personality assessments which contain positive feedback (Reber 1985). The Pollyanna mechanism is a continuation of this, whereby there is a tendency to accept positive comments as true and reject negative ones (Reber 1985). This is also influenced by positive feedback and reinforcement.

French et al (1998) assessed the accuracy and applicability of generalised horoscopes and personal horoscopes. In their three condition experiment, the Barnum statement was rated as more accurate than false or personal horoscope for all conditions. They demonstrated that people felt the "generalised horoscope statement", which had been created by the experimenters, was a more accurate assessment than the participant's own star sign character statement as recorded for that month. This faith in the Barnum statement was linked to the strength of belief in the accuracy horoscopes generally (table 1).

MEANS FOR ACCURACY RATING (1-5)

	BARNUM HOROSCOPE	FALSE HOROSCOPE	GENUINE HOROSCOPE
STRONG BELIEF IN HOROSCOPES (n=7)	4.14	3.29	3.29
MODERATE BELIEF (n=31)	3.77	3.00	3.16
NO BELIEF (n=14)	3.29	3.07	2.79
WHOLE GROUP	3.69	3.06	3.08

(After French et al 1998)

Table 1 - Results from French et al (1998).

This research (1) used an attitude questionnaire testing for belief in spirituality. Spirituality includes spiritualism which refers to a "belief that departed spirits communicate with and show themselves to the living, especially through mediums" and "systems or doctrines founded on this" (Buckley 1982 p1023). In this attitude questionnaire, a general belief in spirituality was used, which incorporated such aspects as

general religious faith, predestination and belief in astrology.

The basis for this study was that people who are susceptible to the Barnum effect are more open to suggestion in terms of spiritual awareness. The research hypothesis predicted a positive correlation between personal accuracy scores on a character description and scores on a spiritual awareness questionnaire. The character assessment was linked to the Myers-Briggs Type Indicator (MBTI) (Myers and Myers 1977) personality traits.

METHOD

A pilot study of the attitude questionnaire on spiritual awareness used thirty statements on a five-point Likert rating scale. Item analysis of differences of 3-4 points reduced the statements to eleven. Table 2 gives the statements used.

1. Fate determines the future
2. There is such a thing as destiny
3. Reincarnation is possible
4. Life is predetermined
5. There is a benevolent force
6. Everyone has a soulmate
7. Guardian angels and ghosts exist
8. Fate can alter the direction of life
9. Life after death exists
10. There is no life after death
11. There is a logical and scientific explanation for everything

(Items 1-9 are positive; 10-11 negative. Likert scale of 1-5 used; thus total score = 55)

Table 2 - Statements used in spiritual awareness questionnaire.

Twenty adult psychology students were approached. Based upon their MBTI profile, one of sixteen 100-word character statements (Open University 2001) was read out, and the participants were asked to rate its accuracy on a scale of 1 (least) to 10 (most accurate). This score was correlated with the score on the spiritual awareness questionnaire (maximum score = 55). Ethical guidelines were followed at all times.

RESULTS

The character assessment gained a mean of 7.5 out of 10 in accuracy (range = 7), with two participants giving a score of ten (table 3). The mean score for the

spiritual awareness questionnaire was 32 from a possible maximum 55 (range = 34) (table 4).

The Spearman rank coefficient was 0.31, but it was not significant at the 5% value (critical value = 0.38; N = 20; Greene 1990).

ACCURACY SCORE	NUMBER OF PARTICIPANTS
3	2
6	2
7	5
8	5
9	4
10	2

Table 3 - Distribution of scores for accuracy scores of character statements.

QUESTIONNAIRE SCORE	NUMBER OF PARTICIPANTS
11-21	1
22-32	10
33-43	8
44-55	1

Table 4 - Distribution of scores of spiritual awareness questionnaire.

DISCUSSION

The positive correlation between the spiritual awareness questionnaire and the perceived accuracy of MBTI generalised character statements was not significant at the 5% level.

The main problem was that the spiritual awareness questionnaire was too general on spiritual issues. Future research would need a more focused questionnaire. It may also be useful to use the same character statement rather than sixteen different ones (2).

FOOTNOTES

1. This research was carried out as part of the DSE202 Summer School at the Open University in 2001. There were four researchers involved.

2. The MBTI distinguishes sixteen personality types (Open University 2001).

REFERENCES

Buckley, E (1982) The Concise Oxford Dictionary Oxford: Oxford University Press

French, C et al (1998) Belief in astrology: a test of the Barnum effect. In Frazier, K (ed) Encounters with the Paranormal

Greene, J (1990) DSE202 Methodological Handbook: Part 1 Milton Keynes: Open University

Myers, P & Myers, K (1977) Myers-Briggs Type Indicator Palo Alto, CA; Consulting Psychologists Press

Open University (2001) DSE202 Summer School Tutor's Handbook Milton Keynes: Open University

Reber, A (1985) Dictionary of Psychology (2nd ed) Harmondsworth: Penguin

Heidi Reinicke

Article written September 2001

THE SOCIAL CONSTRUCTION OF SERIAL MURDER: SOME THOUGHTS

The "cold-blooded senseless" serial killer holds incredible fascination in the Western society today. The endless images in the media present a stereotype of the typical serial killer as a "brutal, blood-thirsty monster" (Hickey 1997).

For those individuals working in the fields of catching or understanding serial killers, it is important to have a clear definition of "serial murder" (the preferred term). Hickey (1986) includes all premeditated killings of three or more victims over time, rather than on one particular occasion (mass murderer). Most importantly, there will be a linkage of common factors among the victims - eg: the type of victim chosen.

In the USA, there appears to be an increasing number of serial murders. For example, the media picked up on the figure of 5000 victims per year from a US Senate hearing with the FBI in the 1980s. This figure is believed to be an overestimation (Hickey 1997).

In one of the best assessments of serial murders in the USA, Hickey (1997) records 399 serial murderers responsible for between 2526-3860 homicides between 1800-1995. But 303 of these include serial murderers who have been active between 1961-1995 (table 1). Up to 1970, there was an average of 0.9 offenders per year, and after 1970, there were 9.9. Clearly there has been a dramatic increase in the number of serial murderers in recent years.

TIME PERIOD	NUMBER OF SERIAL MURDERERS
1800-1820	3
1821-1840	4
1841-1860	3
1861-1880	8
1881-1900	5
1901-1920	16
1921-1940	26
1941-1960	32
1961-1980	173
1981-1995	129
TOTAL	399

(After Hickey 1997)

Table 1 - Number of serial killers in USA by time period, 1800-1995.

It is important to note that the increasing number recently could be due to increased apprehension by the police, and/or better statistical recording of offences. If the increase is "real", rather than a statistical anomaly, "the odds of becoming a victim are minuscule when one considers the size of the population as a whole" (Hickey 1997).

Serial murderers seem to be a phenomena of the USA. Hickey (1997) mentions 75 serial murderers outside the USA since 1800. This is not a systematic figure, and again many crimes could go unreported.

Let us assume that serial murder is a crime that is most common in the USA. What factors in the US society could account for this? I want to suggest that the extensive media interest in recent years has "encouraged" serial murder. Within modern "consumer capitalism" (Brewer 2001), as in the USA, "fame" is a commodity that can be "purchased" and is a sign of status. "Fame" involves television crews making films about you. But "fame" does not necessarily have to be for something "constructive".

Many serial murderers have a background that is ordinary in the sense of having no particular skills to become famous, or even noticeable. They may do a menial job. For example, Arthur Shawcross, who killed eleven women in 1988-89, cut food for salads in a diner.

In a society of "salvation by fame", individuals are "driven" to find ways to achieve "fame". To be apprehended for a series of gruesome killings leads to attention and "fame" (Brewer 2002).

Hickey (1997) notes 54 "Hollywood films" with a storyline about serial murderers between 1990-5, compared to 20 in the 1970s and four in the 1950s (table 2) (1). These are fictional portrayals as well as "real-life documentaries" on television and non-fictional accounts of serial murderers. The media interest is often with "high body counts" and the minutiae of the killings. The media does not have to be accurate in their portrayal of serial murderers, but it is "fame" to have a film made or book written about you.

The explanations for why individuals commit serial murders are complex, and no simple cause will suffice. But the existence of such excessively high media interest, and "fame" must contribute to the reasons for the behaviour.

TIME PERIOD	NUMBER OF "HOLLYWOOD FILMS"
1920-9	2
1930-9	3
1940-9	3
1950-9	4
1960-9	12
1970-9	20
1980-9	23
1990-5	54
TOTAL	121

(After Hickey 1997)

Table 2 - Number of "Hollywood films" about multiple-homicides.

FOOTNOTE

1. Interestingly, "serial killer" is not mentioned that often in the brief description of films in "Satellite TV" in the UK in September 2002. On the satellite film channels, there were 1709 films in that month, six included "serial killer" in the brief description about the film.

On all the channels (satellite and terrestrial in England) in the weeks commencing the 7th and 14th September 2002 (based on "TV and Satellite Weekly" magazine), there were nine references to "serial killer" in details about the programmes (one programme being a documentary, the rest, films). But many more films were about the topic of multiple-homicide.

REFERENCES

Brewer, K (2001) Normality, consumerism, and civil religion Orsett Psychological Review 2, 13-30

Brewer, K (2002) Social hierarchies, status, and the social construction of the self Orsett Psychological Review 6, 2-7

Hickey, E (1986) The female serial murderer Journal of Police and Criminal Psychology 2, 2, 72-81

Hickey, E (1997) Serial Murderers and Their Victims (2nd ed) Belmont, CA: Wadsworth

Kevin Brewer

Article written February 2003